JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 43, 261–272 (1973)

Nonlinear Programming Applied to State-Constrained Optimization Problems

D. O. NORRIS

Department of Mathematics, Ohio University, Athens, Ohio 45701

Submitted by George Leitmann

Necessary conditions for the minimization of a differentiable function subject to differentiable equality and inequality constraints are given. These conditions are applied to a control problem with a state-variable inequality constraint. The connection with earlier work of Neustadt and Jacobson, Lele, and Speyer is given.

1. INTRODUCTION

The purpose of this paper is twofold. First, necessary conditions for the minimization of a differentiable function subject to differentiable equality and inequality constraints are given. These conditions are applicable to problems which may be formulated in infinite-dimensional spaces. Specifically, let X, Y, and Z denote normed linear spaces and let $f: X \to R^1$, $g: X \to Y$, and $h: X \to Z$ denote Frechet differentiable functions. Suppose Z contains a closed convex cone B such that $B^0 \neq \varnothing$ and B has its vertex at the origin. Let $N(g) = \{x: g(x) = 0\}$. (This notation will be used henceforth to designate the null set of a function.) Necessary conditions are obtained for the following nonlinear programming problem:

Minimize f on $N(g) \cap h^{-1}(B)$.

Much work has been done on this problem and problems related to it. The chief advantage to the work presented here is that the necessary conditions are relatively easy to apply. The reader should consult [1]–[3] for references to and discussion of related work.

Recently, Jacobson, Lele, and Speyer [3] introduced new necessary conditions for the optimality of state-constrained problems. They presented a proof of the necessary conditions based upon formally adjoining constraints by use of Lagrange multipliers. The authors suggested that a rigorous proof based upon Neustadt's general theory of extremals [4] should be possible. The second purpose of this paper is to show that a rigorous proof can be obtained by direct application of the nonlinear programming results obtained here. In addition, it will be shown that a special case of Neustadt's results for the state-constrained problem can also be obtained.

It should be noted here that the necessary conditions obtained in Section 3 of this paper, while adequate to handle the problem treated in [3], are not as general as the conditions obtained in [4], for example. The chief drawback is the requirement that the contraints be defined by differentiable functions. This restriction makes it possible to obtain necessary conditions which are relatively easy to apply. It should also be noted that the necessary conditions obtained in [4] only apply to problems for which the equality constraints are finite dimensional. The Lagrange multiplier rule obtained in Section 2 does not suffer from this restriction so that a class of infinite-dimensional problems can be treated by the methods described here.

2. A Lagrange Multiplier Rule

Several results and concepts from [1] and [5] are needed in order to present the Lagrange multiplier rule for minimizing f on $N(g) \cap h^{-1}(B)$. In the interests of clarity the pertinent results will be summarized here. It should be noted that while the nonlinear programming problem posed here is not the most general possible, it is adequate to handle most optimal control problems in which the state equations are given as difference equations, ordinary or partial differential equations, or integral equations. In most control problems the state and control constraints are defined by functional relationships and for such problems the theory developed here applies. Furthermore, the conditions which need to be checked in order to apply the necessary conditions seem to be relatively easy to verify particularly when g is affine. More general formulations in which some of the constraints appear as set constraints give rise to necessary conditions which are difficult to apply (e.g., see [6] and [7]).

THEOREM 2.1. *Let X and Y denote normed linear spaces. Let $g\colon X \to Y$ denote a continuous affine transformation and let $F\colon X \to R^1$ have a nonzero derivative, $DF(x_0)$ at x_0. If F has a critical point on $N(g)$ at x_0 (i.e., $DF(x_0)\,\omega = 0$ for all $\omega \in N(g) - x_0$), then there is a nonzero $y' \in Y'$ (the algebraic dual of Y) such that*

$$DF(x_0)\,\omega + y' Dg(x_0)\,\omega = 0 \qquad \textit{for all } \omega \in X.$$

This is a partial statement of Theorem 3.4 given in [1]. It is shown in [1] that it is not possible, in general, to have y' continuous. However, in many applications it will happen that y' is continuous and then the usual representation theorems may be applied to represent y'. For instance, if $Dg(x_0)(N(DF(x_0)))$ is closed, then y' is continuous (see [1, Theorem 3.1]), or, if there is an $h_0 \notin N(DF(x_0))$ such that $Dg(x_0)\, h_0 \notin \overline{Dg(x_0)\, N(DF(x_0))}$, then y' is continuous. If $Dg(x_0)$ is onto then $Dg(x_0)\,(N(DF(x_0))$ is closed and y' is continuous. Thus, the classical multiplier rule is a special case of Theorem 2.1 in the affine case.

The multiplier rule to be presented here is a simple consequence of Theorem 2.1 above and the following result due to Nagahisa and Sakawa [5].

THEOREM 2.2. *Let X and Z denote normed linear spaces. Let $f: X \to R^1$ and $h: X \to Z$ denote differentiable functions. Let $A \subset X$. Let $B \subset Z$ denote a closed convex cone with vertex at the origin and nonempty interior. If f has a minimum on $A \cap h^{-1}(B)$ at x_0, then there are $\eta^* \in R^1$ and $z^* \in Z^*$ (the topological dual of Z), not both zero, such that*

(i) $\eta^* Df(x_0)\, \omega + z^* Dh(x_0)\, \omega \geqq 0$

for all $\omega \in K$, where K is a closed convex cone with vertex zero contained in the tangent cone $TC(A, x_0)$ of A at x_0 (i.e., $\omega \in TC(A, x_0)$ if and only if there are sequences $\{x_n\} \subset A$, $\{\lambda_n \geqslant 0\} \subset R^1$ such that $x_n \to x_0$ and $\lambda_n(x_n - x_0) \to \omega$).

(ii) $z^* h(x_0) = 0$.

(iii) $\eta^* \geqslant 0$ *and* $z^* z \leqslant 0$ *for all* $z \in B$.

The Nagahisa–Sakawa theorem and Theorem 2.1 can now be applied to the case where $A = N(g)$ for $g: X \to Y$, and the main result of this section follows.

THEOREM 2.3. *Let X, Y, and Z denote normed linear spaces. Let $f: X \to R^1$, $g: X \to Y$, and $h: X \to Z$ denote differentiable functions. Let $B \subset Z$ denote a closed convex cone with nonempty interior and vertex at zero. Suppose x_0 is a regular point of $N(g)$ (i.e., $TC(N(g), x_0) = N(Dg(x_0))$).*

If f has a minimum on $N(g) \cap h^{-1}(B)$ at x_0, then there are $\eta^* \in R^1$, $z^* \in Z^*$, not both zero, and $y' \in Y'$ such that

(i) $\eta^* Df(x_0)\, \omega + z^* Dh(x_0)\, \omega + y' Dg(x_0)\, \omega = 0$ *for all* $\omega \in X$,

(ii) $z^* h(x_0) = 0$, *and*

(iii) $\eta^* \geqslant 0$ *and* $z^* z \leqslant 0$ *for all* $z \in B$.

Proof. Let $F(x) = \eta^* f(x) + z^* h(x)$, where η^* and z^* given in Theorem 2.2 and $A = N(g)$. From Theorem 2.2 it follows that $DF(x_0)\, \omega \geqslant 0$, where $\omega \in K$, K a closed convex cone contained in $TC(N(g), x_0)$. But, x_0 is a regular point of $N(g)$ so that $TC(N(g), x_0) = N(Dg(x_0))$ and therefore K may be chosen as $N(Dg(x_0))$. In this case, since $N(Dg(x_0))$ is a subspace it follows that $DF(x_0)\,\omega = 0$ on $N(Dg(x_0))$, i.e., F has a critical point on $N(Dg(x_0))$. If $DF(x_0) \neq 0$, then by Theorem 2.1 there is a nonzero $y' \in Y'$ such that (i) $\eta^* Df(x_0)\,\omega + z^* Dh(x_0)\,\omega + y' Dg(x_0)\,\omega = 0$ for all $\omega \in X$. Parts (ii) and (iii) are immediate consequences of Theorem 2.2. If $DF(x_0) = 0$, then $\eta^* Df(x_0) + z^* Dh(x_0) = 0$ and $y' = 0'$ will give the desired result.

It should be noted that if g is affine, then every point of $N(g)$ is regular. Thus, the application of the multiplier rule to linear control problems is very easy provided that the defining functions are differentiable. It has been shown by Flett [8] that if $Df(x_0)$ is surjective and g has a continuous derivative, then $x_0 \in N(g)$ is a regular point. In optimal control problems this condition (i.e., $Df(x_0)$ surjective) holds if it is assumed that the adjoint equations are completely controllable.

3. State-constrained Control Problems

Consider the following problem:

Minimize $P(x_u(T))$

subject to the constraints

(i) $x_u'(t) = f(x_u(t), u(t), t),\ x_u(0) = x_0$;

(ii) $Q(x_u(T)) = 0$;

(iii) $S(x_u(t), t) \leqslant 0,\ t \in [0, T]$;

where $P\colon R^n \to R^1$, $f\colon R^{n+r+1} \to R^n$, $Q\colon R^n \to R^q$, $S\colon R^{n+1} \to R^s$, and $u\colon R^1 \to R^r$, such that P, f, Q, and S have continuous derivatives. $S(x, t) \leqslant 0$ means that each component of $S(x, t)$ satisfies the inequality. Hereafter, this problem will be refered to as Problem (SC).

In [3], Jacobson, Lele, and Speyer treat a special case of this problem. In their treatment $x' = f(x, u)$, $S(x_u(t)) \leqslant 0$, and both u and S are scalar valued. Neustadt [4] treats a more general problem than Problem (SC). In his treatment f is assumed to have a continuous partial in x only, control constraints are allowed, Q is allowed to involve the initial states, but S is scalar valued. Problem (SC) or variants of it have been treated by many other authors and [2] or [3] should be consulted for further references. It will be shown in this section that a special case of Neustadt's results can be obtained by application of the necessary conditions. It will also follow that the Jacobson–Lele–Spyer necessary conditions are a special case of the necessary conditions for Problem (SC).

No control constraints have been incorporated in the formulation of Problem (SC). In any specific problem in which the control constraints can be given in the form of $R(u, t) \leqslant 0$, there is no conceptual difficulty in obtaining necessary conditions. In fact, the space Z can be chosen to be a product of two spaces—one for $S(x, t)$ and one for $R(u, t)$. The reason a constraint of the form $R(u, t) \leqslant 0$ is not included here is that such a constraint can assume a variety of forms such as $\int_0^T \| u(t)\|^2 \, dt \leqslant 1$ or $| u_j(t)| \leqslant 1$, $j = 1, 2, \ldots, r$, and, consequently, the space defining the constraint $R(u, t) \leqslant 0$ depends upon the form of the constraint.

Although Problem (SC) is a special case of the example presented by Neustadt, it should be noted that Theorem 2.3 is not a special case of Neustadt's general theory because equality constraints are allowed to be in an infinite-dimensional space.

In order to apply Theorem 2.3 to Problem (SC) the following identifications are made. Let $X = L_\infty(I, R^r)$ denote the space of R^r-valued, essentially bounded, measurable functions defined on $I = [0, T]$ with norm

$$\| u \| = \max_{1 \leqslant j \leqslant r} \operatorname*{ess\,sup}_{t \in I} | u_j(t)|;$$

let $Y = R^q$, and let $Z = C(I, R^s)$, where $C(I, R^s)$ denotes the space of R^s-valued continuous functions defined on I with norm

$$\| z \| = \max_{1 \leqslant j \leqslant s} \sup_{t \in I} | z_j(t)| \, .$$

Note that $B = \{z \in C(I, R^s)\colon z_j(t) \leqslant 0, j = 1, 2, \ldots, s\}$ is a closed convex cone in Z such that B has a nonempty interior and vertex at zero.

Define $F\colon L_\infty(I, R^r) \to R^1$ such that $F(u) = P(x_u(T))$. Define $g\colon L_\infty(I, R^r) \to R^q$ such that $g(u) = Q(x_u(T))$. Define $h\colon L_\infty(I, R^r) \to C(I, R^s)$ such that $[h(u)]\,(t) = S(x_u(t), t)$. Problem (SC) can now be formulated as the nonlinear programming problem treated in Theorem 2.3, viz., minimize F on $N(g) \cap h^{-1}(B)$. In order to compute the derivatives DF, Dg, and Dh, the derivative Dx_u of x_u with respect to u is needed. Note that

$$x_u\colon L_\infty(I, R^r) \to C(I, R^n),$$

hence Dx_u is a linear operator with the same range and domain. Consequently, Dx_u will be represented as a function defined on $I \times I$.

THEOREM 3.1. *Let $x_u\colon L_\infty(I, R^r) \to C(I, R^n)$ be given by*

$$x_u(t) = x_0 + \int_0^t f(x_u(s), u(s), s) \, ds,$$

where $f\colon R^{n+r+1} \to R^n$ has continuous derivatives. Then Dx_u may be represented by

$$\Psi(t, s) \, D_2 f(x_u(s), u(s), s),$$

where D_i denotes the partial derivative of f with respect to the ith *variable and Ψ is the fundamental matrix solution of*

$$a'(t) = D_1 f(x_u(t), u(t), t)\, a(t).$$

Proof. Dx_u exists provided there is a linear operator

$$L: L_\infty(I, R^r) \to C(I, R^n)$$

such that

$$x_{u+h} - x_u - Lh = o(h).$$

Claim $Lh = a_h$, where a_h is the unique solution to

$$a'(t) = D_1 f(x_u(t), u(t), t)\, a(t) + D_2 f(x_u(t), u(t), t)\, h(t), \qquad a(0) = 0.$$

Clearly L is linear in h because

$$a_h(t) = \int_0^t \Psi(t, s)\, D_2 f(x_u(s), u(s), s)\, h(s)\, ds.$$

Now,

$$\begin{aligned} x_{u+h}(t) - x_u(t) - a_h(t) = \int_0^t \{ & f(x_{u+h}(s), u(s) + h(s), s) - f(x_u(s), u(s), s) \\ & - D_1 f(s)\, a_h(s) - D_2 f(s)\, h(s)\}\, ds, \end{aligned}$$

where $D_i f(s) = D_i f(x_u(x), u(s), s)$. Since f has a continuous derivative it follows that

$$\begin{aligned} x_{u+h}(t) - x_u(t) - a_h(t) = \int_0^t \{ & D_1 f(s)\, [x_{u+h}(s) - x_u(s)] + D_2 f(s)\, h(s) \\ & + \omega(s, h(s)) - D_1 f(s)\, a_h(s) - D_2 f(s)\, h(s)\}\, ds, \end{aligned}$$

where

$$\lim_{\|h(s)\| \to 0} \frac{\|\omega(s, h(s))\|}{\|h(s)\|} = 0$$

uniformly in s.

For convenience choose

$$\|h(s)\| = \max_{1 \leqslant j \leqslant r} |h_j(s)|.$$

Then

$$x_{u+h}(t) - x_u(t) - a_h(t) = \int_0^t \{D_1 f(s)\, [x_{u+h}(s) - x_u(s) - a_h(s)] + \omega(s, h(s))\}\, ds,$$

and it follows with $b(t) = x_{u+h}(t) - x_u(t) - a_h(t)$ that

$$\max_{1\leqslant j\leqslant n} | b_j(t)| \leqslant \int_0^t K \max_{1\leqslant j\leqslant n} | b_j(s)| \, ds + \int_0^t \max_{1\leqslant j\leqslant n} | \omega_j(s, h(s))| \, ds,$$

where K is a uniform bound on $\| D_1 f(s)\|$. Thus,

$$\| b(t)\| \leqslant \int_0^t K \| b(s)\| \, ds + \int_0^T \| \omega(s, h(s)\| \, ds.$$

First assume $\int_0^T \| \omega(s, h(s)))\| \, ds \neq 0$. Then by Gronwall's lemma,

$$\| b(t)\| \leqslant \int_0^T \| \omega(s, h(s))\| \, ds \exp \int_0^T K \| b(s)\| \, ds.$$

Since $b \in C(I, R^n)$,

$$\| b \| = \sup_{t\in I} \max_{1\leqslant j\leqslant n} | b_j(t)| \qquad \text{and} \qquad \| b \| \leqslant M \int_0^T \| \omega(s, h(s)\| \, ds.$$

Therefore, with

$$\| h \| = \operatorname*{ess\,sup}_{t\in I} \max_{1\leqslant j\leqslant r} | h_j(t)| \, ,$$

let $\epsilon > 0$ be given. Then there is a $\delta > 0$ such that $\| h \| < \delta$ implies

$$\begin{aligned} \frac{\| b \|}{\| h \|} = \frac{\| x_{u+h} - x_u - a_h \|}{\| h \|} &\leqslant \frac{M \int_0^T \| \omega(s, h(s))\| \, ds}{\| h \|} \\ &\leqslant \frac{M \int_0^T \epsilon \| h(s)\| \, ds}{\| h \|} \leqslant M\epsilon T, \end{aligned}$$

so that $x_{u+h} - x_u - a_h = o(h)$. If $\int_0^T \| \omega(s, h(s))\| \, ds = 0$ for some $h \neq 0$, then $\| b(t)\| \leqslant \int_0^t K \| b(s)\| \, ds$ and it follows that

$$\| b(t)\| = \| x_{u+h}(t) - x_u(t) - a_h(t)\| = 0$$

for all t so that

$$x_{u+h} - x_u - a_h = 0 = o(h). \qquad \text{Q.E.D.}$$

Now the necessary conditions for Problem (SC) can be derived. They will appear in an integral form. In corollaries to the next theorem alternate forms of the necessary conditions will be given.

THEOREM 3.2. *Suppose* $\mathrm{u} \in L_\infty(I, R^r)$ *is an optimal solution to Problem* (SC). *In addition, assume that the range of* $\int_0^T \Psi(T, s) \, D_2 f(x_{\mathrm{u}}(s), \mathrm{u}(s), s) \, h(s) \, ds$ *is* *all of* R^n *(i.e.,* $a' = D_1 fa + D_2 fh$ *is completely controllable) and that* $DQ(x_u(T))$ *has full rank.*

A necessary condition that u be optimal is that

$$\lambda(t)\, D_2 f(x_u(t), u(t), t) = 0,$$

a.e. on I, where

$$\lambda(t) = [\eta_0 DP(x_u(T)) + y' DQ(x_u(T))]\, \Psi(T, t) + \int_t^T d\eta^*(s)\, D_1 S(x_u(s), s)\, \Psi(s, t),$$

and where each component of the s-dimensional row vector η^* *is nondecreasing on* $[0, T]$. *Furthermore,* η_j^* *is constant on intervals for which the* jth *component* $S(x_u(t), t) < 0$ *and* η_j^* *is nondecreasing when the* jth *component of* $S(x_u(t), t) = 0$.

Proof. With F, g, and h defined as given in the discussion preceding Theorem 3.1, it follows by application of Theorem 2.3, the hypothesis of complete controllability, and the hypothesis that $DQ(x_u(T))$ have full rank that u is a regular point and that there are $\eta_0 \geqslant 0$, $y' \in R^q$ and $z^* \in C(I, R^s)^*$, such that

$$\eta_0 DF(u)\, \omega + y' Dg(u)\, \omega + z^* Dh(u)\, \omega = 0 \qquad \text{for all } \omega \in L_\infty(I, R^r).$$

Apply the chain rule to F, g, and h to obtain

$$\eta_0 DP(x_u(T))\, Dx_u(T)\, \omega + y' DQ(x_u(T))\, Dx_u(T)\, \omega + z^* D_1 S(x_u(t), t)\, Dx_u(t)\, \omega = 0$$

for all $\omega \in L_\infty(I, R^r)$. From Theorem 3.1 and the fact that

$$\eta_0 DP(x_u(T))\, Dx_u(T) + y' DQ(x_u(T))\, Dx_u(T) + z^* D_1 S(x_u(t), t)\, Dx_u(t)$$

represents an element in $L_\infty(I, R^r)^*$, it follows that

$$[\eta_0 DP(x_u(T)) + y' DQ(x_u(T))] \int_0^T \Psi(T, s)\, D_2 f(s)\, \omega(s)\, ds + \int_0^T d\eta^*(t)\, D_1 S(x_u(t), t) \int_0^t \Psi(t, s)\, D_2 f(s)\, \omega(s)\, ds = 0$$

for all $\omega \in L_\infty(I, R^r)$, where $\eta^* \in NBV(I, R^s)$ represents $z^* \in C(I, R^s)^*$, i.e., η^* is an R^s-valued normalized function of bounded variation.

Now $D_1 S(x_u(t), t)$ and $\Psi(t, s)$ are continuous, hence $D_1 S(x_u(t), t)\, \Psi(t, s)$ is product measurable. Furthermore, since $D_2 f(s)\, \omega(s)$ is measurable, it follows that $K(t, s) = D_2 f(s)\, \omega(s)$ is product measurable. Thus, $D_1 S(x_u(t), t)\, \Psi(t, s)\, D_2 f(s)\, \omega(s)$ is product measurable. Finally, since

$D_1S(x_u(t), t)\,\Psi(t, s)\,D_2f(s)\,\omega(s)$ is essentially bounded, it follows that $D_1S(x_u(t), t)\,\Psi(t, s)\,D_2f(s)\,\omega(s)$ is integrable. Then by Fubini's theorem,

$$\int_0^T d\eta^*(t)\,D_1S(x_u(t), t)\int_0^t \Psi(t, s)\,D_2f(s)\,\omega(s)\,ds$$
$$= \int_0^T ds \int_s^T d\eta^*(t)\,D_1S(x_u(t), t)\,\Psi(t, s)\,D_2f(s)\,\omega(s)$$

and the necessary condition may be written as

$$\int_0^T ds\,\Big\{[\eta_0 DP(x_u(T)) + y'DQ(x_u(T))]\,\Psi(T, s) + \int_s^T d\eta^*(t)\,D_1S(x_u(t), t)\,\Psi(t, s)\Big\}\,D_2f(s)\,\omega(s) = 0$$

for all $\omega \in L_\infty(I, R^r)$. Therefore, it must happen that

$$\Big\{[\eta_0 DP(x_u(T)) + y'DQ(x_u(T))]\,\Psi(T, s) + \int_s^T d\eta^*(t)\,D_1S(x_u(t), t)\,\Psi(t, s)\Big\}\,D_2f(s)$$
$$= \lambda(s)\,D_2f(x_u(s), u(s), s) = 0, \qquad \text{a.e. on } I.$$

From conditions (ii) and (iii) of Theorem 2.3 it follows that $\eta_0 \geqslant 0$, each component η_j^* of η^* is non-decreasing such that

$$\lim_{t\to 0^+} \eta_j^*(t) = 0, \qquad \text{and} \qquad \int_0^T d\eta^*(t)\,S(x_u(t), t)\,dt = 0.$$

Now $S(x_u(t), t) \leqslant 0$ and is continuous in t, hence if a component $S_j(x_u(t'), t') < 0$ for some $t' \in (0, T)$, then $S_j(x_u(t), t) < 0$ for all t in an open interval (t_i^j, t_{i+1}^j). Thus, $[0, T]$ is the union of a collection of open intervals on which $S_j(x_u(t), t) < 0$ and a collection of closed intervals or degenerate closed intervals (i.e., points) on which $S_j(x_u(t), t) = 0$. In order for $\int_0^T d\eta^*(t)\,S(x_u(t), t) = 0$ it follows that η_j^* is constant on each open interval on which $S_j(x_u(t), t) < 0$ and η_j^* is nondecreasing elsewhere. Q.E.D.

COROLLARY 3.1. *If* u *is continuous, then* $\lambda(t)\,D_2f(x_u(t), u(t), t) = 0$ *everywhere on* I.

Proof. λ is right continuous and D_2f is continuous since u is continuous, hence λD_2f is right continuous. Then, by the standard persistence of sign argument with minor modifications it follows from $\int_0^T \lambda(t)\,D_2f(t)\,\omega(t)\,dt = 0$ for all $\omega \in L_\infty(I, R^r)$ that $\lambda(t)\,D_2f(x_u(t), u(t), t) = 0$ everywhere on I.

Q.E.D.

COROLLARY 3.2. *If $s = 1$ (i.e., S is a scalar-valued constraint),* u *is continuous, and $D_1S(x_{\mathrm{u}}(t), t)\, D_2 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t) \neq 0$ for all t such that $S(x_{\mathrm{u}}(t), t) = 0$, then η^* can have no jumps.*

Proof. Suppose $S(x_{\mathrm{u}}(t_1), t_1) = 0$ and η^* has a jump at t_1. Then

$$\lambda(t_1^+) = \lambda(t_1^-) - D_1S(x_{\mathrm{u}}(t_1), t_1)\, [\eta^*(t_1^+) - \eta^*(t_1^-)].$$

Now, $\lambda(t_1^+) = \lambda(t_1)$ since λ is right continuous and $\lambda(t_1)\, D_2 f(t_1) = 0$. Furthermore

$$\lambda(t_1^-)\, D_2 f(t_1^-) = \lim_{t \to t_1^-} \lambda(t)\, D_2 f(t) = 0$$

so it can be concluded that

$$D_1S(x_{\mathrm{u}}(t_1), t_1)\, D_2 f(x_{\mathrm{u}}(t_1), \mathrm{u}(t_1), t_1) = 0$$

contrary to hypothesis. Q.E.D.

This result is a special case of the result stated in [3, Theorem 6]. The argument above suffices to show that [3, Theorem 6] holds for $p = 1$. The condition $D_1S(x_{\mathrm{u}}(t), t)\, D_2 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t) \neq 0$ on $S(x_{\mathrm{u}}(t), t) = 0$, is often refered to as a regularity condition (see, for example, [9]).

In the next corollary, the necessary conditions given in Theorem 3.2 are cast into the form given by Neustadt. There is a minor difference in that η used here is nondecreasing and $\eta(T) = 0$, whereas in Neustadt's work it is taken to be nonincreasing and $\eta(T) = 0$. This difference is due to the fact that in the derivation of the necessary conditions here, it was assumed that $z^*z \leqslant 0$ on B, whereas Neustadt uses the reverse inequality. Furthermore, since differentiability in control and no control constraints have been assumed, the necessary conditions do not involve a maximum over the control set. As pointed out earlier, there is no difficulty in incorporating an explicit control constraint given in the form $R(u, t) \leqslant 0$.

COROLLARY 3.3. *Suppose S has a continuous second derivative. Then necessary conditions for* u *to define a minimum for Problem* (SC) *are*

(i) $[\tilde{\psi}(t) - \eta(t)\, D_1S(x_{\mathrm{u}}(t), t)]\, D_2 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t) = 0$, *a.e., on I, where*

(ii) $\tilde{\psi}'(t) = -\tilde{\psi}(t)\, D_1 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t) + \eta(t)\, D_1 p(x_{\mathrm{u}}(t), t)$,

(iii) $p(x, t) = D_1S(x, t) f(x, \mathrm{u}(t), t) + D_2S(x, t)$, *and*

(iv) $\tilde{\psi}(T) = \eta_0 DP(x_{\mathrm{u}}(T)) + y' DQ(x_{\mathrm{u}}(T))$.

Proof. Define $\eta(t) = \eta^*(t) - \eta^*(T)$, where η^* is given in Theorem 3.2. Define $\tilde{\psi}(t) = \lambda(t) + \eta(t)\, D_1S(x_{\mathrm{u}}(t), t)$ where λ is given in Theorem 3.2. Then (i) is immediate from Theorem 3.2.

Let

$$\psi(t) = [\eta_0 DP(x_{\mathrm{u}}(T)) + y' DQ(x_{\mathrm{u}}(T))]\, \Psi(T, t)$$

and let

$$\psi_1(t, s) = D_1 S(x_{\mathrm{u}}(s), s)\, \Psi(s, t).$$

Claim

$$\tilde{\psi}(t) = \psi(t) - \int_t^T \eta(s) \frac{\partial}{\partial s} \psi_1(t, s)\, ds.$$

Integrate $\int_t^T d\eta(s)\, \psi_1(t, s)$ by parts to obtain

$$\int_t^T d\eta(s)\, \psi_1(t, s) = -\, \eta(t)\, D_1 S(x_{\mathrm{u}}(t), t) - \int_t^T \eta(s) \frac{\partial}{\partial s} \psi_1(t, s)\, ds.$$

Since $\lambda(t) = \psi(t) + \int_t^T d\eta(s)\, \psi_1(t, s)$ the desired result follows.

Now, differentiate $\tilde{\psi}(t) = \psi(t) - \int_t^T \eta(s)\, (\partial/\partial s)\, \psi_1(t, s)\, ds$ and make use of the fact that

$$\frac{\partial}{\partial t} \Psi(s, t) = -\, \Psi(s, t)\, D_1 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t)$$

to obtain (ii) and (iii). Part (iv) follows by direct substitution. Q.E.D.

Finally, the necessary conditions of Jacobson, Lele, and Speyer [3] can be obtained. The author cannot justify the formal operations used in [3] to verify that η^* has a piecewise continuous derivative. Consequently, it will be assumed that η^* has the desired property.

COROLLARY 3.4. *Suppose* η^* *given in Theorem* 3.2 *has a piecewise continuous derivative, then necessary conditions for* u *to define a minimum for Problem* (SC) *are*

(i) $\lambda(t)\, D_2 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t) = 0$, *a.e., on* I,

(ii) $\lambda'(t) = -\, \lambda(t)\, D_1 f(x_{\mathrm{u}}(t), \mathrm{u}(t), t) - \eta^{*\prime}(t)\, D_1 S(x_{\mathrm{u}}(t), t)$ *a.e. on* I,

and

(iii) $\lambda(T) = \eta_0 DP(x_{\mathrm{u}}(T)) + y' DQ(x_{\mathrm{u}}(T))$.

Proof. Since η^* is piecewise differentiable it follows that

$$\lambda(t) = [\eta_0 DP(x_{\mathrm{u}}(T)) + y' DQ(x_{\mathrm{u}}(T))]\, \Psi(T, t) + \int_t^T \eta^{*\prime}(s)\, D_1 S(x_{\mathrm{u}}(s), s)\, \Psi(s, t)\, ds + j(t)$$

where $j(t)$ is a step function. Then λ' exists except where j has jumps and is given by

$$\lambda'(t) = -[\eta_0 DP(x_u(T)) + y' DQ(x_u(T))]\, \Psi(T, t)\, D_1 f(t) - \eta^{*\prime}(t)\, D_1 S(x_u(t), t)$$
$$- \int_t^T \eta^{*\prime}(s)\, D_1 S(x_u(s), s)\, \Psi(s, t)\, ds D_1 f(t)$$
$$= -\lambda(t)\, D_1 f(t) - \eta^{*\prime}(t)\, D_1 S(x_u(t), t) \qquad \text{a.e. on } I.$$
$$\lambda(T) = \eta_0 DP(x_u(T)) + y' DQ(x_u(T))$$

follows by direct substitution. Q.E.D.

References

1. D. O. Norris, A generalized Lagrange multiplier rule for equality constraints in normed linear spaces, *SIAM J. Control* **9** (1971), 561–567.
2. L. W. Neustadt, An abstract variational theory with applications to a broad class of optimization problems. II. Applications, *SIAM J. Control* **5** (1967), 90–137.
3. D. H. Jacobson, M. M. Lele, and J. L. Speyer, New necessary conditions of optimality for control problems with state-variable inequality constraints, *J. Math. Anal. Appl.* **35** (1971), 255–284.
4. L. W. Neustadt, A general theory of extremals, *J. Comput. System Sci.* **3** (1969), 57–92.
5. Y. Nagahisa and Y. Sakawa, Nonlinear programming in Banach Spaces, *J. Optimiz. Theory Appl.* **4** (1969), 182–190.
6. P. P. Varaiya, Nonlinear programming in Banach space, *SIAM J. Appl. Math.* **15** (1967), 284–293.
7. M. Guignard, Generalized Kuhn–Tucker conditions for mathematical programming problems in a Banach space, *SIAM J. Control* **7** (1969), 232–241.
8. T. M. Flett, On differentiation in normed vector spaces, *J. London Math. Soc.* **42** (1967), 523–533.
9. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, "The Mathematical Theory of Optimal Processes," Wiley, New York, 1962.